EPA Information Strategy Stakeholders Meeting

March 8-9, 2001

Facilitator's Flip Chart Notes

Key Issues/Concerns

- Occurrence data
- How to deal with developing info systems – capture and delivery to public
- Data drives policy
- Implementation of new rules
  - big task
  - data quality
- Integration of data – utility, state and federal levels
- Integration of facility ID information
- Keeping data set simple enough to be able to work with it
- Data collection and reporting burden – potential impacts on data quality
- Data quality
- Reporting burden
- Transition to data management
- Data integrity throughout decision making process.
- How to upload data
- How accurate data reporting
- SDWIS/FED development – listening for needs and ideas
- Selection of quality metrics
- Harmonization of data
- Effects on listed chemicals
- Ideas and options for improving information systems
- Public access to data regarding their water and enforcement data
- Implications for small business
- Implications for implementation of SDWIS-State
- Good quality data (occurrence and exposure)
- Data completeness for rule implementation
- Support for UCMR
- Understanding clear objective

Questions Regarding Existing Systems

- Will SDWARS be available to public?
  - *ultimately, yes*
- Which systems intended to be primary for public access?
  - *ultimately, SDWIS-Plus*
- Will state labs be able to upload to SDWARS?

- *lab that does analysis will do the reporting*
- *hopefully via www*

- Will there be opportunity for QA/QC by utilities in phase 2?
  - *yes*

I. a) Source Water Contamination Prevention

- Guidance to states on TMDL – information/data?
  - *with stakeholders input – yes*
- Intend to look at CWA information to see if it is relevant to drinking water – if so, will look at it - if needed for EPA national picture – otherwise stay at state level -- geo ref (at intakes and wells) information needed to establish this possible relationship
- Will publish strategy in 3-6 weeks
- Drinking water information being used for CWA assessments – not seen in positive light - need more/better communication between SDWA and CWA people at EPA HQ and Regions
  - *this strategy is intended to start addressing this*
- Will more information be required from states for this?
  - *4 data items – in SDWS*
  - *no regulatory requirement to provide the information but a need for the information*
- Don't count widgets – focus on information needed to assess results
  - lot of counting inherent in table presented
  - locals not currently mandated to report this information to state – his effort required to get this data
  - *key to assess whether the information supports national picture, id who collects and how to get it to systems*
- Concern – source water prevention - a PWS responsibility?

  Systems only one piece of local architecture of prevention for systems

I. a) OGWDW Data Needs

- Need to ask: what analyses will be done, and how will they be done? (what questions and how data will be used to answer them)
- What documentation required to demonstrate quality of what is done with the data to ensure quality
- Can't see how to link UIC to PWS information without GIS
- Different places in different states have data on PWS, UIC and source water protection
- Caution against establishing an unwarranted relationship
- Relate all to geographic location – using consistent location information – link all different types of data to location. Then can ID potential for relationships – need good geospatial data
- Where is responsibility for collecting source water protection data –
  - voluntary
  - not required of PWS
  - state cup is full – collection of locational data means something else won't get done
- Real tension when creating central integrated system – means changes/compromises for managers and users of existing systems

- Those responsible for collecting the data should have primary weight
- Lat/long might answer data integration issues – then user needs to be able to identify their lat/long to make use of data (for intakes and wells)
- We have mechanism to get geospatial information for every PWS
- EPA information needs:
  - for six year review and other auxiliary needs
  - measures effectiveness of protection programs
  - identify contaminants for regulation
  - identify steps to provide source water protection
- Make best use of existing data collection programs (e.g. USGS)
- States have similar data needs themselves
- Locational data valuable, but knowing the limits of that data equally valuable – caution regarding using it inappropriately
  - must designate appropriate use/interpretation of locational data
- Also prepared to drop data requirements that are not needed
- Need to assess effectiveness of SDWA/ What is best way to do this? May not be in information systems we are looking it
- Threshold questions:
  1- need same level/extent of information from all PWSs?
  2- do we need to know about every violation?
- Have to ask what our job is in order to answer threshold questions
- Should we be doing business differently from when we set this up years ago (mostly to support enforcement and compliance)?
- Don't gear information collection requirements towards bad actors – trust PWS summary data
- Ontario – reviewing regulations and information required – and changed regulations – hold PWSs accountable to report their information
- Federal agency has to decide what its business is and the information needed to do it
- Different state programs have different standards – often more stringent than Federal – so collecting violation data based on state standards will be different than based on national MCL (and different/comparable to different states)
- Intelligent reporting software can tell you if lat/long matches a zip code, etc. – can help address burden issue
- Primacy program sets up fundamental tension between feds and state (butting heads) especially when it relates to a national DW program.
- Organize around customers/suppliers
- What at source of state/federal "butting heads"?
  - Requests for more data perceived as threat or display of lack of trust – for use in enforcement and compliance
  - Communication problem – e.g. performance partnership agreements
  - Tension because EPA management talks to state management but different communication at level of programs
  - Money from feds does not go to data reporting
  - Need consistent messages:
    Legislature

EPA management
EPA programs
State management
State programs
PWSs

- If EPA willing to make real change – get together with states (current approach doesn't work well)
- What data states provide to EPA really helps to protect public health?

➢ *EPA responsible to oversee and back stages state programs – needs this data to do this – see patterns and analysis of performance of program and take enforcement action in e.g. regions cases*
  - *Is this what it should be?*
  - Information on compliance doesn't do this
  - Provide data so EPA oversight is credible
  - Data currently asked for does not do this (not a valid use of the data being collected – especially lat/log; treatment technology.)

- If counting beans all the time, not paying as much attention to other things
- Utility responsible to <u>customer</u>
- One EPA business need – identifying <u>emerging</u> contaminants (UCMR)

<u>I. b) What data Needed to Evaluate Success and Where Can it From:</u>

- Parametric data from states may not address the data needs
- Reg. to inform public <u>plus</u> additional requirements – e.g. CCR
- For program evaluation collect through mid-year review (regions)
  - EPA come review the documents
  - Or 3-year review
- <u>Either</u> parametric or violation data
- Need occurrence with parametric

<u>EPA National Drinking Water Program</u>:

- Oversight and backstopping
  - compliance data critical for this

- Inform Public

- Address emerging contaminants
- Evaluate existing regulations
- Developing new regulations
  - affordability/economics of treatment
  - co-occurrence and treatment technologies
- Sensitive sub-pops

- Source water protection

- For backstop/oversight
  - violations/compliance data from states (outcome focus)
  - others systems set up for collecting other program evaluation needs
  - don't ask every system, but a sampling
- Need to define/clarify questions before identifying information needs
- Need basics:
  - what is in water, where
  - characteristics of system/effectiveness of treatment

- EPA Question:
- *Is it the reporting burden or that you don't want EPA to have it (the data)?*
  - ➢ Concern regarding how data will be used, interpreted by EPA or public and the limits of the data:
    - some of it is wrong
    - data collected for are reason may have different quality control that data collected for another purpose
    - needs to be properly handled
  - ➢ No problem sharing SDWIS/State data once up and running
  - ➢ Data – fine metadata if people willing to go get it
  - ➢ ICR huge burden – resistance not about sharing data
  - ➢ Concern regarding second-guessing state data/program

- Need for reporting tracking:
- distinguish between monitoring or reporting violation versus an MCL violation
  - lots more work!

I. c) Parametric Data

- Need data system with place holder for this data – don't want a new database
- Are we talking about voluntary, required, regulatory reporting?
- *not yet decided – what is appropriate and representative?*
  - Go through state
  - Any such requirement inappropriate as a regulation going around primacy state (conflict between primacy – direct implementation
  - UCMR not part of primacy – since by def unreg some states pleased not to be involved
  - Some states pleased NOT to be involved
- Are some data elements only needed for some period of time (i.e. effectiveness of treatment)?
- This data used especially for six-year review – to see if regs are doing what they are intended to do.
- Need meta–data on sample by sample basis – for accurate comparability of data
- Labs provide analytical data – their ID is included with their reports.
- Per system basis for list I contaminants (all large and some small systems) $3-4000/system (1 yr. monitoring)

I. e) How Communicate Better Between EPA and States

- Involve data people from states more in regulation development (*why have they not been there?* – has not been the emphasis)
- Participate in ASDWA meetings – and get data on the agenda – or establish sub-group on data
- Data management folks have taken a back seat – need to step up
- AWWA source water – ask what data elements are needed
- Include OEI in more of the discussions
- Hard for data and technology people to talk to each other

- tend to work separately – need to hear each other (lesson from ICR)
  - fine points get missed if technology people not involved – IT people need to be involved early
- Data sharing issues are management/policy issues – not for data "handlers" (data management staff)
  - how to collect data requested
  - data management a small piece
- EPA IMB newsletter is good
- Internally, EPA not getting their people together
- Drivers at state level (TX) are responsibilities to state legislature and to EPA and to water systems (what state agrees to do for a fee) (these become "performance measures" and have implications throughout the system)
- EPA sometimes doesn't research their own data before asking for more
  - smarter use of existing data
- Key is understanding the questions and data needed and then how collect, etc. (and if it needs to be collected or already exists)

I. f)

- No additional data requirements beyond PWSS program
- Geospatial data – very complex
  - vectors
  - adjacent/branches of watershed
  - etc.
  - need to be very clear what you want/need
  - very expensive
- Source water assessment data may be collected already, but in different systems – question of getting it to other systems
  - *link to this data*
- DW program is changing – prevention is emerging as more important – need to identify data needs and availability
- Look to waste-water program also – link WW and DW to share data
- Interactive mapping projects
- Need to know hydro geologic sensitivity of water source with source water assessment data
- States have history of protection activities that these new requirements interfere with.
- Secondary users of data – will use data in ways or for purposes not originally intended
- measuring success of protection

- no outbreak = success
- can only really "measure" failure
- outbreak is a trailing indicator – we need to look for leading indicators

• Need to define the value of data requested: define the business need – does the data meet the need?
  - EPA not the core business managers, but need to evaluate success of the business (of states)
  - EPA does get the requests for data – from other agencies/congress/public, etc. not always anticipated in past

II. a)

• How much data is electronically submitted?
  - most or all to EPA
  - to states: 0-80% - varies a lot (more chem. and rads than microbial)
  - electronic reporting makes burden much more manageable - need to get smart software to those reporting

  Why:
  - large labs can report electronically
  - small systems – report on paper (whoever owns the data reports)
  - capability of multiple labs varies
  - slow process – quality control issues – state wants oversight of quality control
  - small labs may not have LIM

• Goal – make electronic data exchange part of lab certification requirement
• Guidance for state reporting should be part of role development and rule when promulgated – not after the fact. (also, don't leave specific data requirements for after the rule).
• Proposed shift:
  - EPA define core data requirements
  - Establish partnerships/utilities to agree on other valuable information; state will collect and save.

- Develop directory of who has what data
- *(EPA – this is where we want to go)*

• Create a template for small systems (Ontario model)
• Pennsylvania:
  - going paperless
  - existing pilot projects
  - improve data quality – 1 set of data
  - inventory – SDWIS plus much more

II. b)

• Depends on questions you're asking
• Directory approach – not pull all data into one place

- Can CCR be modified to address these needs? A good vehicle that already exists

II. c)

- Desired frequency of new releases?
  - SDWIS/FED – annually (or not at all)
  - SDWIS/State – more often
    (currently other way around)
- No substitute for one-on-one communication with states (e-mail, phone)
- Go towards web-bases XML (post data outside firewall) – more efficient
- Still need violations linkages
  - caution against limiting flexibility of 4 methods – has implications for states
- More simultaneous release of SDWIS-Fed and State.

III.

- Leave SDWIS/FED as is and focus resources on developing new approach (XML/CDX)
  - takes states long time to put modifications in place
- equiv. question for states – don't' want to invest in addressing SDWIS/FED if something new is coming
- Question: how deal with SDWIS/FED regarding rules in the pipeline?
  - Just because SDWIS Fed can't handle data for new rule doesn't mean states are not collecting the required information
  - Not ask states to submit to 2 systems
  - How can EPA get this if SDWIS/FED not updated
    - States can provide spreadsheets
    - EPA provide format (what fits in SDWIS Feds in it and what doesn't – until new system is ready)
- In interim: create new modules that can link to new system
- Stop SDWIS/FED.....add appropriate fields to SDWIS/State and use SDWIS/State for now (to provide for needs of SDWIS/FED users)
- Keep as much as you can of what works well enough (data transfer causes problems so minimize it)
- Keep old system going as is -- modular approach to building new ones – transfer information module by module.
- Try not to change table structure but can add new codes to existing fields
- Continue to keep SDWIS/FED up to date (include modifications) – until new system can accommodate needs.

IV. Data Quality

- Do not require use of SDWIS/State..... a lot invested by some states in other systems
- Show us that data is really needed to answer questions
- Under-reporting of violations not in data issues – determination by state regarding appropriate reporting/monitoring frequency
- Only require major violations reported

- Problem: most violations at small systems – harder to get the data – emphasis on small systems needed
  - Community. vs. non-community esp. (for violations data)
- Unique identifier is critical
- Eliminate problem of recording violations for 30 chems because 1 sample missed
- Data quality not that bad – methodology to evaluate data quality paints states in worst light
- Difficulty in accessing systems to do QA/QC is a serious problem – fixing this would help (utility perspective) – standard internet access
- Timeliness related to above
- Provide cushion on late reporting (not call it a violation right away)
- Make it easier to update violations

Outstanding Questions

- How does stakeholder involvement process for this interface with stakeholder process for source water protection strategy development?
  - SWP – separate stakeholder process to identify measures – then hand off to this process